Solution 1: Risk Minimizers for the Log-Loss

(a)

$$
\begin{aligned}
\pi_c^* &= \arg\min_{c\in[0,1]} \mathbb{E}_{xy}\left[L\left(y,c\right)\right] = \arg\min_{c} \mathbb{E}_{y}\left[L\left(y,c\right)\right] \\
&= \arg\min_{c} \mathbb{E}_{y}\left[-y\log\left(c\right) - (1-y)\log\left(1-c\right)\right] \\
&= \arg\min_{c} -\log\left(c\right) \underbrace{\mathbb{E}_{y}\left[y\right]}_{=\mathbb{P}(y=1)=\pi} - \log\left(1-c\right)\underbrace{\mathbb{E}_{y}\left[1-y\right]}_{=1-\pi} \\
&= \arg\min_{c} -\left[\pi\log\left(c\right) + (1-\pi)\log\left(1-c\right)\right]
\end{aligned}
$$

Taking the derivative with respect to c and setting it to 0:

$$
\begin{aligned}
&\Rightarrow \frac{\partial}{\partial c}\left[-\pi\log\left(c\right) + (1-\pi)\log\left(1-c\right)\right] \overset{!}{=} 0 \\
&\Rightarrow -\frac{\pi}{c} + \frac{1-\pi}{1-c} = 0 \\
&\Rightarrow c(1-\pi) \qquad = (1-c)\pi \\
&\Rightarrow c \qquad = \pi \\
&\Rightarrow \pi_c^* \qquad = \mathbb{P}(y=1)
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathcal{R}_l(\pi_c^*) &= \mathbb{E}_{xy}\left[L\left(y,\pi\right)\right] \\
&= \mathbb{E}_{y}\left[-y\log\left(\pi\right) - (1-y)\log\left(1-\pi\right)\right] \\
&= -\pi\log(\pi) - (1-\pi)log(1-\pi) \\
&= H(y) \ \ (= \text{Entropy!})
\end{aligned}
$$

(c) $\hat{\boldsymbol{\theta}}$, the optimal constant model in terms of the *empirical* risk, is given by $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \mathcal{R}_{emp}(\boldsymbol{\theta})$.

$$
\begin{aligned}
\mathcal{R}_{emp}(\boldsymbol{\theta}) &= \sum_{i=1}^{n} L\left(y^{(i)}, f\left(\mathbf{x}^{(i)}\right)\right) \\
&= \sum_{i=1}^{n} \log\left(1+\exp(-y^{(i)}\boldsymbol{\theta})\right)
\end{aligned}
$$

As $L(y,\boldsymbol{\theta}) = log\left(1+\exp(-y^{(i)}\boldsymbol{\theta})\right)$. Taking the derivative:

$$
\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\theta}}\mathcal{R}_{emp}(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \frac{1}{1+\exp(-y^{(i)}\boldsymbol{\theta})}\left(+\exp(-y^{(i)}\boldsymbol{\theta})\right)(-y^{(i)}) \\
&= -\sum_{i=1}^{n} y^{(i)}\frac{\exp(c)}{1+\exp(c)} \\
&= -\sum_{y^{(i)}=1}^{n} (1)\frac{\exp(-\boldsymbol{\theta})}{1+\exp(-\boldsymbol{\theta})} - \sum_{y^{(i)}=-1}^{n} (-1)\frac{\exp(\boldsymbol{\theta})}{1+\exp(\boldsymbol{\theta})} \\
&\overset{!}{=} 0
\end{aligned}
$$

This is equivalent to:

$$\begin{aligned}
\sum_{y^{(i)}=-1}^{n} \frac{\exp(\boldsymbol{\theta})}{1+\exp(\boldsymbol{\theta})} &= \sum_{y^{(i)}=1}^{n} \frac{\exp(-\boldsymbol{\theta})}{1+\exp(-\boldsymbol{\theta})} \\
n_{-}\frac{\exp(\boldsymbol{\theta})}{1+\exp(\boldsymbol{\theta})} &= n_{+}\frac{1}{1+\exp(\boldsymbol{\theta})} \\
\frac{n_{+}}{n_{-}} &= \exp(\boldsymbol{\theta}) \\
\boldsymbol{\theta} &= \log(\frac{n_{+}}{n_{-}})
\end{aligned}$$

Solution 2: Risk Minimizers for the Brier Score

(a) The loss using Brier score is given by $L\left(y, \pi(\mathbf{x})\right) = \left(y - \pi(\mathbf{x})\right)^2$.

$$\begin{aligned}
\pi_c^* &= \arg\min_c \mathbb{E}_{xy}\left[L\left(y, c\right)\right] \\
&= \arg\min_c \mathbb{E}_{y}\left[(y-c)^2\right] \\
&= \arg\min_c \mathbb{E}_{y}\left[y^2 - 2yc + c^2\right] \\
&= \arg\min_c \mathsf{Var}_y(y) + \pi^2 - 2c\pi + c^2 \\
&= \arg\min_c \pi(1-\pi^2) + \pi^2 - 2c\pi + c^2 \\
&= \arg\min_c \pi - \pi^2 + \pi^2 - 2c\pi + c^2 \\
&= \arg\min_c \pi - 2c\pi + c^2 \\
&= \arg\min_c -2c\pi + c^2
\end{aligned}$$

Where we used $\mathsf{Var}(y) = \mathbb{E}(y^2) - \left[\mathbb{E}(y)\right]^2$.
Taking the deriviative with respect to c and setting it to 0:

$$\begin{aligned}
&\Rightarrow \frac{\partial}{\partial c}\left[-2c\pi + c^2\right] \overset{!}{=} 0 \\
&\Rightarrow -2\pi + 2c = 0 \\
&\Rightarrow \pi \qquad = c \\
&\Rightarrow \pi_c^* \qquad = \mathbb{P}(y=1)
\end{aligned}$$

(b)

$$\begin{aligned}
\mathcal{R}_l(\pi_c^*) &= \mathbb{E}_{xy}\left[L\left(y, \pi\right)\right] \\
&= \mathbb{E}_{y}\left[(y-\pi)\right] \\
&= \mathbb{E}_{y}\left[y^2 - 2y\pi + \pi^2\right] \\
&= \pi - 2\pi^2 + \pi^2 \\
&= \pi - \pi^2 \\
&= \pi(1-\pi) \\
&= \mathsf{Var}_y(y)
\end{aligned}$$

Loss	**Risk Minimizer**	**Bayes Risk**	**Optimal Constant Model**	**Bayes Risk (Optimal Model)**
L2	$\mathbb{E}_{y\|\mathbf{x}}(y\|\mathbf{x}) = f^*(\mathbf{x})$	$\mathcal{R}^*_{L2} = \mathbb{E}_x[\mathsf{Var}_{y\|x}(y\|x)]$	$\mathbb{E}_y[y] = f^*_c$	$\mathsf{Var}_y(y) = \mathcal{R}_{L2}(f^*_c)$
0/1	$h^*(\mathbf{x}) = \arg\max_{C\in\mathcal{Y}} \mathbb{P}(y = C\|\mathbf{x} = \mathbf{x})$	$\mathcal{R}^*_{0/1} = 1 - \mathbb{E}_x[\max_{C\in\mathcal{Y}} \mathbb{P}(y = C\|\mathbf{x} = \mathbf{x})]$	Exercise 2	Exercise 2
Log	$\pi^*(\mathbf{x}) = \mathbb{P}(y = 1\|\mathbf{x} = \mathbf{x})$	$\mathcal{R}^*_l = \mathbb{E}_x[\mathrm{H}_{y\|x}(y\|x)]$ exp. cond. entropy (chapter 13)	$\pi^*_c = \mathbb{P}(y = 1)$	$\mathrm{H}_y(y) = \mathcal{R}_l(\pi^*_c)$
Brier	$\pi^*(\mathbf{x}) = \mathbb{P}(y = 1\|\mathbf{x} = \mathbf{x})$	$\mathcal{R}^*_B = \mathbb{E}_x[\mathsf{Var}_{y\|x}(y\|x)]$ $(= \mathcal{R}^*_{L2})$	$\pi^*_c = \mathbb{P}(y = 1)$	$\mathsf{Var}_y(y) = \mathcal{R}_B(\pi^*_c)$